Osteoarthritis and Cartilage (2009) **17,** 579–585

doi:10.1016/j.joca.2008.10.004

Osteoarthritis and Cartilage

International Cartilage Repair Society

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Reproducibility of computer-assisted joint alignment measurement in OA knee radiographs

A. K. O. Wong†*, D. Inglis‡, K. A. Beattie†, A. Doan†, G. Ioannidis†, J. Obeid§, J. D. Adachi|| and A. Papaioannou†

† *Department of Medicine, McMaster University, Hamilton, ON, Canada*
‡ *Department of Civil Engineering, McMaster University, Hamilton, ON, Canada*
§ *Department of Kinesiology, McMaster University, Hamilton, ON, Canada*
|| *Department of Medicine, McMaster University, St. Joseph's Healthcare, Hamilton, ON, Canada*

Summary

Objectives: (1) To investigate the reproducibility of computer-assisted measurements of knee alignment angle (KA) from digitized radiographs of osteoarthritis (OA) participants requiring total knee arthroplasty (TKA) and (2) to determine whether landmark choice affects the precision of KA measurements on radiographs.

Methods: Using a custom algorithm, femoral, central, and tibial measurement-guiding rules were interactively placed on digitized posteroanterior fixed-flexion knee radiographs by mouse control and positioned according to different anatomic landmarks. The angle subtended by lines connecting these guiding rules was measured by three readers to assess interobserver, intraobserver and experience–inexperience reproducibility. Test–retest reproducibility was evaluated with duplicate radiographs from a healthy cohort. Reproducibility was assessed using root-mean square coefficients of variation (RMSCV%). The Bland–Altman method was performed on data obtained from varying anatomic landmarks (confidence interval, CI = 95%).

Results: From 16 healthy and 30 TKA participants, reproducibility analyses revealed a high degree of intraobserver ($n = 38$, RMSCV = 0.56%), interobserver ($n = 38$, RMSCV = 0.72%), test–retest ($n = 16$, RMSCV = 0.87%) and experience–inexperience ($n = 38$, RMSCV = 0.73%) reproducibility with variances below 1%. Varying the orientation of tibial and femoral rules according to anatomic landmarks produced a difference that exceeded an *a priori* limit of agreement of $-1.11°$ to $+1.67°$.

Conclusion: Our custom-designed software provides a robust method for measuring KAs within digitized knee radiographs. Although test–retest analyses were only performed in a healthy cohort, we anticipate a similar degree of reproducibility in an OA sample. A standardized set of anatomic landmarks employed for KA measurement is recommended since arbitrary selection of landmarks resulted in imprecise KA measurement even with a computer-assisted technique.

Key words: Severe knee osteoarthritis, Total Knee Arthroplasty, Digital, Knee alignment angle, Reproducibility, Anatomic landmarks, Computer-assisted, Severe osteoarthritis.

Introduction

'Knee osteoarthritis (OA) is a multifactorial disease involving structural and functional changes in the joint[1]. Risk factors for knee OA are categorized as constitutional and mechanical, the latter involving factors directly affecting the joint such as muscle weakness, joint injury and possibly malalignment[2]. Although no direct causality has been established for knee malalignment on OA progression, several investigators have suggested such a relationship[3–6].

In patients with knee OA, increased malalignment is frequently associated with advanced disease progression and concomitant functional limitation[6]. The direction of knee alignment deformity is related to either medial or lateral compartment disease[7]. Small changes in joint angle may be indicative of considerable cartilage thinning in either tibiofemoral compartments[3], thus becoming useful for assessing OA progression. In fact, Cicuttini and colleagues demonstrated that a 1° increase in varus angulation is associated with an average annual reduction of 17.7 μL of femoral cartilage[3].

Orthopaedic surgeons confirm post-surgical alignment angle in total knee arthroplasty (TKA) patients to ensure proper installation of the artificial joint. Small changes in knee alignment angle (KA) may reduce the stability of the joint, increase postoperative pain, reduce the longevity of TKA prostheses and possibly predispose patients to further mechanical damage[8]. Considering the costs associated with primary TKA, the longer hospital stay and higher hospital costs associated with revision TKA[9], post-surgical misalignment should be avoided. As such, correct assessment of knee mechanical angle is crucial for improving long-term outcome assessment, and can be facilitated by a technique that allows rapid turnover and feedback to the orthopaedic surgeon.

*Address correspondence and reprint requests to: A. K. O. Wong, Department of Medicine, McMaster University, 501-25 Charlton Ave E, Hamilton, ON, Canada. Tel: 1-905-527-9100; E-mail: wongko@mcmaster.ca
Received 17 June 2008; revision accepted 9 October 2008.

Traditionally, the full-limb radiograph has been required to assess KA. Typically, KA is measured using a method described by Moreland and colleagues[10] in which landmarks are drawn at the hip, knee and ankle. Accuracy of full-limb KA measurement has been questioned due to multiple difficulties encountered in the procedure. An experienced radiologist is required to identify landmarks since it is difficult to discern bone from the dense soft tissue around the hip and ankle. Secondly, multiple X-ray gridded-cassettes must be employed[11]. The feasibility of this method is limited due to the geometric manipulations required to perform the measurement. Further complications arise from lower extremity rotation which results in reduced accuracy and reproducibility[10]. These factors pose an economic disadvantage to full-limb KA assessment and reduce the practicality of the method. While resource-demanding for large scale clinical trials, its feasibility in routine clinical practice is further reduced, thus advocating for a more reliable technique.

In the measurement of KA using full-limb radiography, the 'mechanical' axes of the femur (center of femoral head to center of knee) and tibia (center of knee to center of ankle) are measured. The subtended mechanical angle between the axes represents the angular deviation from the static weight-bearing axis, a line drawn from the femoral head to ankle. The mechanical angle is therefore a true reflection of mechanical load distribution. More recently, Kraus and colleagues developed a new KA assessment technique employing plain knee radiographs[11]. Here, anatomic axes defined by lines bisecting the mid-shaft of the femur and tibia are constructed and the anatomic angle measured is converted back to a mechanical angle by a calibration equation. This conversion factor is important because the anatomic axes only capture the orientation of the femur and tibia at a more distal and proximal site, respectively. While full-limb KA measurement may be more accurate for assessing how angular deformity may affect load distribution, the newer technique appears to be more economical and less tedious. We therefore investigated reproducibility of digital KA measurement in knee radiographs of severe OA participants requiring TKA using a digital method modified from Kraus' validation study[11]. We particularly focused on determining test−retest reproducibility and evaluated whether variation in use of anatomic landmarks affects KA measurement.

Methods

PARTICIPANTS

TKA participants were selected from a previous study analyzing articular cartilage in a pre-arthroplasty population. These participants presenting with either single or bilateral knee OA were recruited through a local orthopaedic surgeon from Hamilton Health Sciences (Henderson Campus) (Hamilton, ON, Canada). Healthy participants were a convenience sample of a previous study analyzing knee minimum joint space width in asymptomatic volunteers[12,13]. All participants had previously consented to have X-rays of their knees acquired. Approval for this study was granted by the Research Ethics Board at McMaster University (Hamilton, ON, Canada) and Hamilton Health Sciences.

RADIOGRAPHY

Posteroanterior (PA) radiographs of the knee were obtained for all participants using the fixed-flexion technique[14]. In the healthy cohort, duplicate radiographs of the non-dominant knee were acquired with half-hour intermission between acquisitions. In the TKA population, single radiographs were acquired of knees recommended for arthroplasty. All examiners who performed KA measurement were blinded to the health status of subjects, and to all KA measurements.

CUSTOM-DESIGNED KA MEASUREMENT SOFTWARE

Knee radiographs were digitized with a radiographic film digitizer (VIDAR® Sierra Plus Digitizer, Herndon, VA) at a bit depth of 12 and 300 dpi resolution. Images were examined using custom-designed software, adjusted according to Kraus' KA measurement protocol[11], allowing mouse-controlled zoom, pan, window-level adjustment, measurement-guiding rules and corresponding distance and angle reading (Fig. 1). Measurement readings were scaled by the software to account for digitizer resolution. All analyses were performed on the same computer workstation.

DEFAULT KA MEASUREMENT

Three measurement-guiding rules: a center rule, a tibial rule, and a femoral rule, were interactively positioned on digitized radiograph by mouse control. By default, end points of the center rule were placed on tips of the tibial spines. The midpoint of this rule was analogous to the center of the knee as illustrated by Moreland *et al.*[10] The femoral rule was adjusted parallel to femoral condyles and the tibial rule parallel to tibial plateau. When all measurement-guiding rules were correctly positioned, their angular alignment was secured by software. A distance of 10.0 cm away from the center rule was automatically prescribed for the tibial and femoral rules. In the case of short radiographs, the tibial and femoral rules were manually dragged to the furthest possible distance. End points of the tibial and femoral rules were positioned on the outer cortical shell. These anatomic landmarks are depicted in

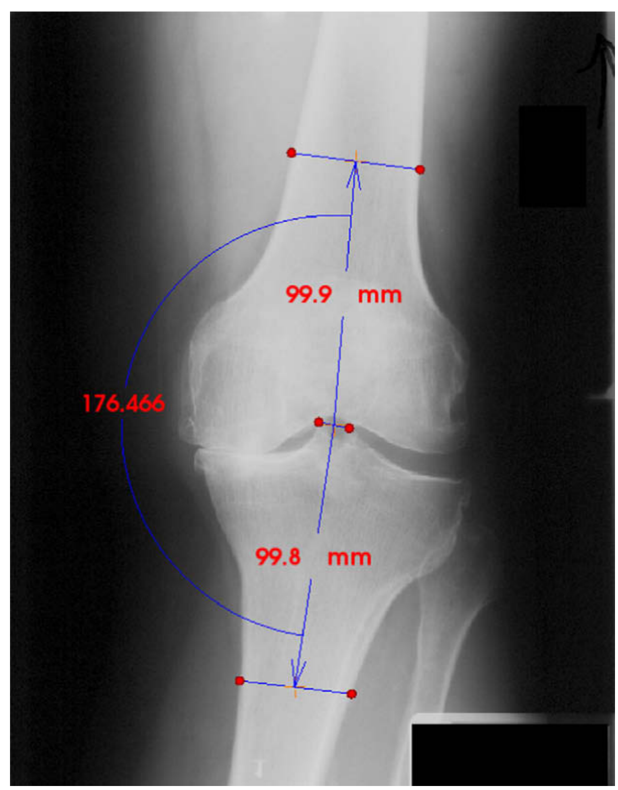

Fig. 1. KA assessment using custom-designed software algorithm under the default set of anatomic landmarks. The center rule in the middle is placed on the tibial spine tips. Both femoral and tibial rules are positioned at 10.0 ± 0.5 cm from the center rule with end points placed on the outer cortical shell and aligned parallel to femoral condyles and tibial plateau, respectively. All guiding rules are bordered by adjustable end points marked in red circles. Femoral and tibial distances from the center rule are indicated in red, as is the subtended anatomic angle. Anatomic angle (θ_A) is separately converted to a mechanical angle (θ_M) according to an equation determined by Kraus and colleagues[11] $\theta_M = (0.69 \times \theta_A) + 53.69$.

Fig. 1. The angle subtended by lines connecting the midpoints of the femoral, center and tibial rules was measured as the anatomic angle (θ_A) and was converted to a mechanical angle (θ_M) according to an equation determined by Kraus[11]:

$$\theta_M = (0.69 \times \theta_A) + 53.69$$

EFFECTS OF VARYING ANATOMIC LANDMARK ON KA MEASUREMENT

To determine whether varying anatomic landmark selection would affect KA measurement, end points and orientations of tibial and femoral rules were adjusted according to various anatomic landmarks: (1) end points of both rules placed on inner vs outer cortical edges, (2) midpoint to midpoint distance from center rule equal to 10.0 ± 0.5 cm vs 5.0–7.0 cm, and (3) femoral rule aligned parallel to femoral condyles vs tibial plateau. Such anatomic landmarks were chosen for their ease of detection by software due to better image contrast and for the purpose of evaluating the robustness of the technique on short radiographs measuring less than 10 cm proximally and distally from the center of the knee. Each of these conditions were varied individually and compared to the aforementioned default conditions for tibial and femoral rule placement. The anatomic landmark protocol was performed by an experienced reader (AKOW).

REPRODUCIBILITY

Two experienced readers (AKOW, KAB) and one inexperienced trained reader (JO) analyzed radiographs from the TKA population. One experienced reader (AKOW) reviewed the duplicate radiographs from the healthy population. KA was measured using the default KA measurement protocol described above. All readers repeated KA measurements on separate days at least 1 week after the first measurement.

DATA ANALYSIS

All mechanical KAs, after conversion from anatomic KAs, were expressed as degrees of deformity away from the quoted average "normal" male KA of 178.5° by Moreland and colleagues[10]. The same normal reference KA was employed by Zhai *et al.*[7] in their male and female cohort for KA assessment. For the purpose of assessing reproducibility and reliability of KA measurement, we have chosen this angle as a normal reference for both genders since a healthy female "normal" angle has not been established in the literature. Hence, negative deformity values represented varus alignment while positive values represented valgus alignment. A Student's one-sample one-sided t test was used to compare 178.5°[10] and global mean KAs for both TKA and healthy participants at the 99% confidence level.

In comparing differences in KA measured between default and different anatomic landmark choice, a Bland–Altman procedure was performed to identify the limit of agreement (LOA) among individual inter-method differences. The LOA is defined as the 95% confidence interval for the difference in measurements between the two methods. Bland–Altman graphs are described by the distribution of individual inter-method differences plotted against the individual inter-method means on the x-axis[15]. Indications of the LOA are marked on the Bland–Altman plot to illustrate the tightness or separation of data. For the purpose of defining how good an LOA we want to achieve, an *a priori* LOA of −1.11° to +1.67° was adapted from Goker and colleagues[16], representing the variability between the manual goniometer method and a digital method (the digital–manual LOA). The digital–manual LOA therefore served as a standard for assessing agreement in digital KA measurement between conditions of varied anatomic landmarks.

Comparisons were made between KA from: two experienced readers on the same day to assess interobserver reproducibility; the same reader on different days for intraobserver reproducibility; and experienced and trained readers on the same day to determine experience–inexperience reproducibility. Test–retest reproducibility was evaluated by comparing measurements obtained on duplicate radiographs for participants in the healthy cohort. All reproducibility analyses were evaluated using root-mean square coefficients of variation (RMSCV%) and root-mean square standard deviations (RMSSD) as previously described[17].

Results

A total of 46 participants were analyzed and comprised of 16 healthy participants (14 female, 2 male; age: 39 ± 12 yr; body mass index (BMI): 24 ± 4 kg m^{-2}) and 30 participants recommended for TKA as a consequence of severe OA (17 female, 13 male; age: 66 ± 9 yr; BMI: 31 ± 6 kg m^{-2}).

Table I

Global assessment of computer-assisted radiographic measures of KAs in TKA participants (n = 38) and healthy cohort (n = 16), expressed as amount of deviation from a 'normal' mechanical angle of 178.5°[10]. Global averages were taken from measurements acquired on all experimental days where the default conditions of anatomic landmark placement were used

	Mean deformity (°)	SD (°)	S.E.M. (°)	P value
TKA participants	−3.18*	5.07	0.34	<0.001
Valgus	4.42*	3.58	0.51	<0.001
Varus	−5.21*	3.11	0.23	<0.001
Healthy cohort	0.73	1.84	0.33	0.043

*Indicates statistically significant difference from healthy alignment angle of 178.5° at the 99% confidence interval.

Table I shows the mean and standard deviations of the KAs measured for the TKA participants and the healthy cohort. Alignment angles in the TKA population ranged from 165.5° to 189.2° with 79% having varus alignment. Meanwhile, the healthy cohort had a tighter KA range: between 175.5° and 182.2°, with 31% varus alignment. Among the TKA participants, both varus and valgus aligned knees ($n = 38$) exhibited mean KAs that significantly deviated from the reported mean "normal" healthy angle of 178.5° (range: 173.5°–182.0°)[10] with 99% confidence level (Table I). The healthy cohort's mean KA ($n = 16$) was significantly different from "normal" at the 95% confidence level (Table I). However, the ranges of KAs in both our healthy cohort and Moreland's reported healthy population were comparable.

EFFECTS OF VARYING ANATOMIC LANDMARK ON KA MEASUREMENT

Although the mean KA measured by the different methods appeared comparable, there was more apparent inter-method imprecision observed on a patient-to-patient basis. When comparing the default conditions vs placing guiding rules at 5.0–7.0 cm, the LOA fell beyond the *a priori* digital–manual LOA of −1.11° to +1.67° by approximately 2-fold on either side of the limit ($n = 38$) (Table II).

In addition, following closely with the digital–manual LOA was the LOA for the default condition vs alignment of the femoral rule parallel to the tibial plateau (Table II). In contrast, placing femoral and tibial rule end points on the inner

Table II

Bland–Altman analysis comparing default conditions for anatomic landmark placement during alignment angle measurement with each of the three varied conditions for anatomic landmark choice [(A) end points of rules placed at edge of inner cortical shells (Inn.Cort.Shell), (B) rules placed 5.0–7.0 cm away from tibial spine midpoint (5.0–7.0 cm), (C) femoral rule oriented parallel to tibial plateau (//Tib.Plat.)]. Analyses were performed only in TKA participants on a single day

	Mean difference (°)	Absolute mean difference (°)	SD of mean difference (°)	95% LOA (°)
Default vs Inn.Cort.Shell	−0.07	0.21	0.26	−0.58–0.43
Default vs 5.0–7.0 cm	0.13	1.14	1.45	−2.70–2.97
Default vs //Tib.Plat.	0.02	0.46	0.65	−1.24–1.29

or outer margin of the cortex seemed to produce little variation in measured KA ($n = 38$) (Table II). This was demonstrated by a narrow LOA and relatively low SD for mean difference. The LOAs are visualized in Fig. 2 for the three comparisons.

REPRODUCIBILITY

Intraobserver, interobserver, experience–inexperience and test–retest reproducibility experiments all yielded RSMCV values below one percent (Table III), though the test–retest RMSCV demonstrated a slightly greater variance than other comparisons. The absolute mean difference between intraobserver measurements for all three readers was in each case, below 1° ($n = 38$). On the other hand, interobserver reproducibility ($n = 38$), experience–inexperience reproducibility ($n = 38$), as well as test–retest reproducibility ($n = 16$) exhibited about twice as large as absolute mean difference. However, these mean differences were still within an acceptable level of reproducibility.

Discussion

We have demonstrated high reproducibility for KA measurement using a computer-assisted method in PA views of knee radiographs. The RMSCVs were well below one percent for all analyses while all reproducibility experiments yielded RMSSDs that were within or just above 1°. Varying anatomic landmark placement during KA measurement revealed considerable inter-method imprecision as illustrated by wide LOAs. To our knowledge, this is the first study to assess test–retest and experience–inexperience reproducibility of a computer-assisted KA measurement technique in plain knee radiographs.

Previous KA study protocols[3,18] were based on work by Moreland *et al.*[10], while others were independently established for full-limb radiographs[16,19,20]. One recent study investigated digital KA measurement in knee radiographs[21]. However, no anatomic landmark use was mentioned, nor was the technique validated. We adapted a KA measurement protocol based on that of Kraus and colleagues who validated manual KA measurement in plain knee radiographs against the traditional full-limb radiograph technique[11]. Our development of a computer-assisted knee radiograph technique also incorporated anatomic landmarks as determined by Moreland *et al.*[10], which allowed for more accurate KA measurements to be made.

Other investigators illustrated a substantial degree of imprecision associated with inconsistent selection of anatomic landmarks for manual KA measurement in full-limb radiographs[16]. In spite of this evidence, there have been no reports of similar studies for digital KA measurement in knee radiographs. Moreover, other investigators have not evaluated test–retest or experience–inexperience precision[16,21], although such information is important for assessing the robustness of a digital technique.

EFFECTS OF VARYING ANATOMIC LANDMARKS ON KA

We encountered considerable patient-to-patient variability in the difference between KA values measured by differential placement of anatomic landmarks. Particularly, the notably wide LOA comparing KA measurements obtained by placing guiding rules at 10.0 cm vs 5.0–7.0 cm away from the tibial spines indicated reduced precision when altering guiding rule distance. This LOA was greater compared to the digital–manual LOA. While the digital–manual LOA of −1.11° to +1.67° appears as a small range, similarly small changes in KA have been proposed to infer progression of OA[16].

Former studies have manually assessed KA at less than 10.0 cm from the tibial spines in short radiographs[7,11]. In our study, there were nine short radiographs wherein placement of anatomic landmarks was restricted within 5.0–7.0 cm. Such a compromised field of view precludes optimal identification of anatomic landmarks and KA values obtained are hence not reliable. Instead, we propose that future KA readings be performed in knee radiographs acquired with sufficient coverage such that measurements can be made at 10.0 ± 0.5 cm both proximally and distally from the tibial spine tips. The 10 cm reference distance was used by Kraus *et al.* for manual measurement of KA in knee radiographs. Since this technique was validated against the full-limb method, it is therefore a more reliable starting point[11]. Prior studies demonstrated greater accuracy in measuring KA at the supracondylar region of the femur rather than in the middle of the shaft, which would otherwise offset the measured mechanical angle[10].

Compared to the wider LOA observed with varying tibial and femoral rule distance, KA measurement was more tolerable to differential alignment of the femoral rule. There do not appear to be any specifications in other investigators' manual or digital techniques describing how the femoral guiding rule (or similar guides) is aligned when determining the medial-to-lateral midpoint of the shaft[10,11,21]. The intuitive approach would be to place guiding rules perpendicular to edges of the cortical bone. However, due to the curvature of the distal femoral neck, this is often not possible. In fact, there would likely be increased intraobserver and interobserver error in determining what is considered perpendicular.

The next obvious landmark to use is the line parallel to the femoral condyles. Moreland and colleagues identified their femoral anatomic axis by connecting the superior aspect of the femoral head to the distal portion of the medial femoral condyle, effectively bisecting the femur mid-shaft[10]. Aligning the femoral guiding rule parallel to the femoral condyles to determine mid-shaft distance was probably a reasonable approximation. In addition, we also aligned the femoral rule to the tibial plateau for the simple reason that it was a more convenient and readily defined landmark to use. Consequently, there would be less room for error. However, we found that the resulting LOA was as large as the digital–manual LOA. With the goal of achieving greater precision and to standardize KA measurement, aligning the femoral rule to the femoral condyles is considered more optimal.

The inconsistent use of anatomic landmarks when measuring KA has led to similarly poor reliability as found in standard clinical readings[22]. In addition, variations in patient positioning have presented challenges to the longitudinal measurement of KA in order to evaluate change over time and this may also be a factor in reduced reproducibility of KA measurements. Positioning devices such as the SynaFlexer™ X-ray positioning frame (Synarc, San Francisco, CA)[14] may aid in improving variation during X-ray acquisition. However, radiographic features in OA knees may inadvertently reduce the ease of measurement. For example, it was often difficult to identify tips of the tibial spines in the presence of osteophytic growth and joint space narrowing (Fig. 3). Ideally, the use of landmarks that are independent of disease characteristics would maximize consistency.

Our study has shown that choice of anatomic landmarks is an important factor to consider when measuring KA.

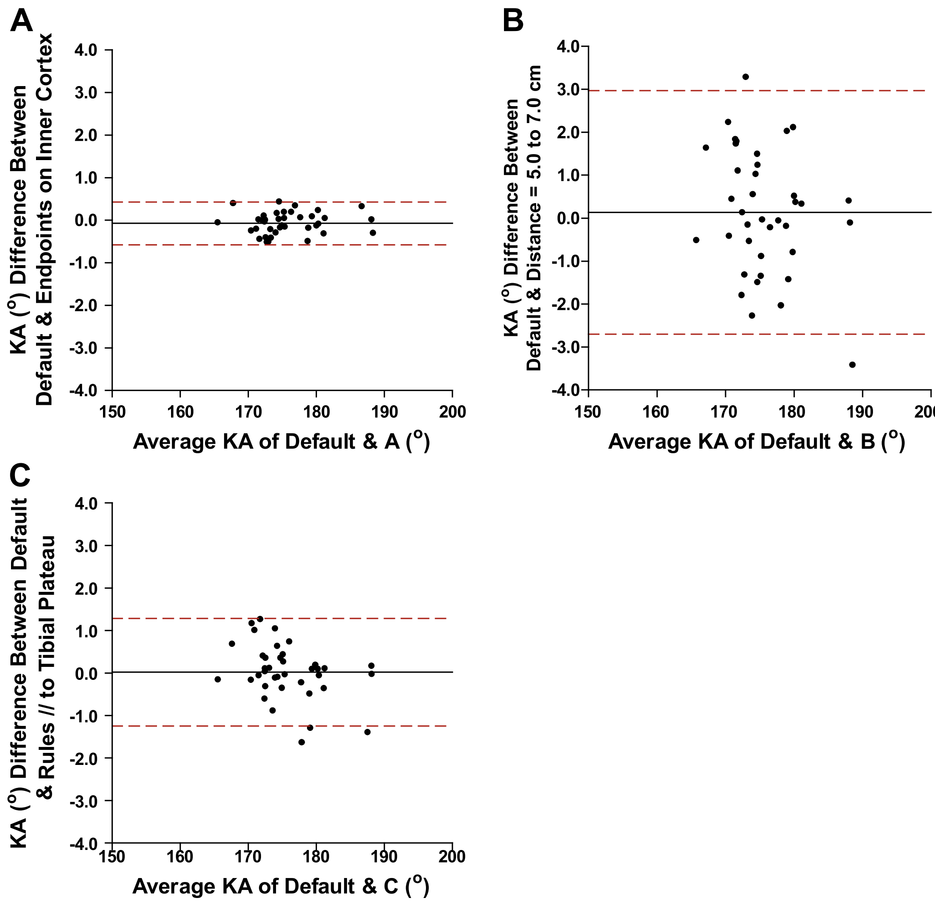

Fig. 2. Bland–Altman plot illustrating the degree of agreement in KA (°) between the default protocol for anatomic landmark placement and (A) with measurement-guiding rules placed on the inner edge of medial and lateral cortical shells, (B) with measurement-guiding rules placed 5.0–7.0 cm away from the midpoint of tibial spines, (C) with femoral rule placed parallel to the tibial plateau. Difference in KAs of individual participants between methods is illustrated along with the global mean difference marked by the solid horizontal line. Upper and lower boundaries representing two standard deviations around the global mean difference is indicated by dashed horizontal lines.

While former studies demonstrating imprecision associated with landmark choice have been reported for manual techniques[22], it would be more important for digital techniques to employ consistent anatomic landmarks for KA measurement. Ultimately, the goal of digital KA assessment is to reduce imprecision and increase efficiency by limiting or eliminating user-intervention.

Table III

Root-mean square measures of coefficient of variation and standard deviation (RMSCV and RMSSD, respectively) for all reproducibility analyses (interobserver, intraobserver, experience–inexperience using TKA participants, and test–retest using healthy cohort). Intraobserver reproducibility measures are expressed as an overall average for all readers ± SD

Reproducibility comparisons	Absolute mean difference (°)	RMSCV (%)	RMSSD (°)
Intraobserver			
Experienced 1 (AKOW)	0.81	0.56	0.99
Experienced 2 (KAB)	0.97	0.83	1.51
Trained (JO)	0.54	0.29	0.51
Average	0.78 ± 0.22	0.56 ± 0.27	1.01 ± 0.50
Interobserver (experienced 1–2)	1.34	0.72	1.27
Exp.–inexp. (experienced 1–trained)	1.34	0.73	1.30
Test–retest (radiograph A–B)	1.67	0.87	1.56

REPRODUCIBILITY OF DIGITAL METHOD

Previous studies employed proprietary software or more widely available open-source image software for measuring KA of OA participants[16]. These software packages have demonstrated a high degree of reproducibility[16,19,21]. The low RMSCV values obtained in our reproducibility experiments indicate that our digital method was capable of producing repeatable KA measurements in OA knees by the same user, different users, inexperienced users and, perhaps most importantly in considering application to longitudinal studies, on separate radiographs of our healthy volunteer knees.

Our absolute mean differences in KA reported for intraobserver measurements being below 1° are considered within the limits of measurement error for conventional manual KA measurement[3]. Particularly, only with precision levels below 1° may clinical KA assessment be more useful since such small changes in alignment may confer considerable biomechanical impacts on the joint over time[3]. Goker and colleagues reported a minimal detectable difference of 0.4° in intraobserver measurements, which was slightly lower than

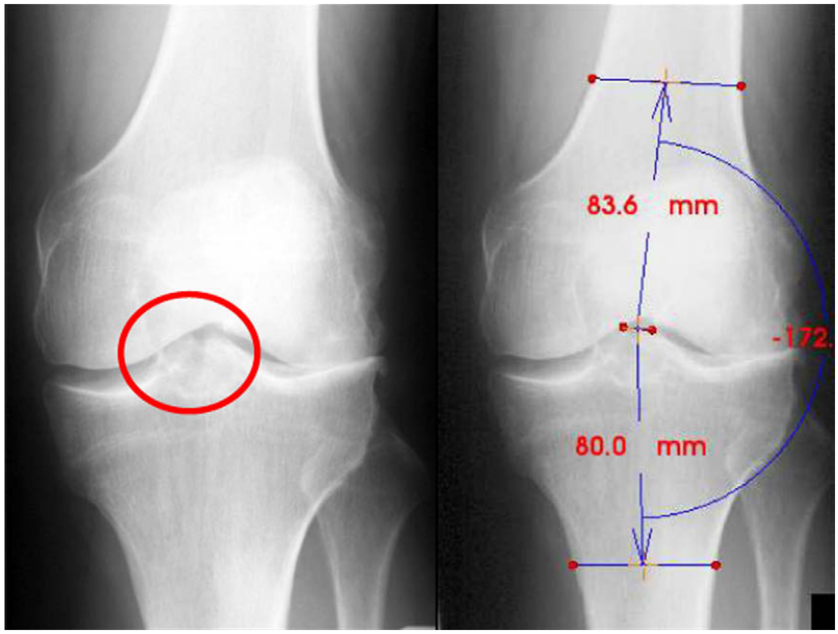

Fig. 3. Osteophytic growth and joint space narrowing preventing clear identification of tibial spine tips and tibial borders for digital knee alignment measurement. Left: medial tibial spine tip unidentifiable behind medial femoral condyle indicated in circle. Right: user-identified landmarks for obstructed tibial spine tips.

our mean difference (Table III). Although the interobserver ($0.38° \pm 1.12°$) and intraobserver ($0.16° \pm 1.12°$) relative mean differences for Takahashi's digital method on knee radiographs were lower, their measurement of KA involved collaboration of several surgeons to identify measurement landmarks[21].

Our interobserver, test–retest and experience–inexperience absolute mean differences were just above 1°. It remains to be determined whether such differences are less than a clinically relevant change over time, at least for the purpose of assessing OA progression. In fact, no longitudinal studies on KA change have been performed to date. Refinement of the software interface and/or improved training of end users may further reduce this minimal detectable difference.

LIMITATIONS

Individual data points that deviated considerably from the mean were shown to be derived from short radiographs, which, as aforementioned, could contribute to reduced precision in KA measurement. Moreover, our test–retest reproducibility experiment was performed in a healthy cohort, which was chosen as a convenience sample as we did not have access to duplicate radiographs in the TKA population. Hence, there are limitations to the interpretation of test–retest reproducibility here reported to be applied against a population of OA participants. Our test–retest absolute mean difference was highest among all reproducibility experiments and this result may be explained by systematic differences in knee positioning during X-ray acquisition and by patient disease characteristics. It is, however, not possible to determine if poorer reproducibility was a result of the actual analysis technique. In fact, we expect that test–retest reproducibility would be poorer if performed in a severe OA cohort due to difficulties in maintaining stability. Nevertheless, the fact that different technologists may have performed the repeat scans might explain some of the variation in measurement.

While the fixed-flexion technique has been well documented[14,23], certain investigators have identified a lack of precision in joint space width measurement[24,25] as a result of variability in tibial rim alignment (TRA)[24,26]. Since joint space width is related to alignment angle, it is not surprising that differential TRA distance may also contribute to test–retest imprecision. Once again, stabilizing patient position may, in fact, result in improved test–retest reproducibility of KA.

Recent editorial comments from Cooke[27] suggest that the use of anatomic axes alone for KA assessment may not be appropriate. Without being able to define the mechanical load bearing axis, the measured KA may not fully represent alignment deformities. While we have not used full-limb radiographs to define the load bearing axis, we did convert our anatomic axes to mechanical axes based on Kraus' former validation study[11]. Also, the mechanical KAs obtained were compared against the 'normal' KA and hence provided the notion of deformity from neutrality.

Meanwhile, we consider that using a 10 cm reference distance in knee radiographs would preclude the occurrence of any distortions proximal and distal to the short axes[27]. However, assuming such distortions remain consistent longitudinally, our technique may be reliable for measuring KA change. Considerations for cost, time, facilities and statistical power would ultimately determine whether digital KA measurement from knee radiographs or full-limb methods would be ideal for clinical assessment in a particular study design[28].

Conclusions

Our high degree of test–retest and experience–inexperience reproducibility demonstrated robustness for digital KA measurement in knee radiographs. Whereas we have only measured test–retest reproducibility in a healthy cohort, we do not anticipate that repeat scans from an OA population would present insurmountable challenges aside from minor hindrance due to joint space narrowing and osteophytosis. Thus, we would expect to see a similar degree of reproducibility in an OA population.

Our results also revealed a considerable degree of imprecision associated with non-standardized anatomic landmark choice when measuring KA using a digital method. Digital methods have been shown by a number of groups to be highly reproducible, at least in single users. We suggest that a standardized set of anatomic landmarks be chosen to improve precision of KA measurement between users. Based on our observations, we propose the following choice of anatomic landmarks for KA measurement: femoral and tibial rules with end points on outer cortical shell, femoral rule aligned with femoral condyles, tibial rule aligned with tibial plateau, and both rules positioned at 10.0 ± 0.5 cm from the midpoint of tibial spine tips.

With more robust software for KA measurement in knee radiographs, practitioners can avoid the more tedious manual goniometer-aided analysis of full-limb radiographs. In effect, the high costs and resources associated with KA assessment can be diminished by using single-cassette X-rays and a computer-assisted method that is simple and reproducible. The increased efficiency of digital KA assessment may ultimately allow readings to be performed in routine clinical practice.

Conflicts of interest

None of the authors have relevant conflicts of interest to declare.

Acknowledgements

We would like to thank our funding from the CIHR Strategic Training Program in Skeletal Health Research (STP 53892).

References

1. Martin JA, Buckwalter JA. Roles of articular cartilage aging and chondrocyte senescence in the pathogenesis of osteoarthritis. Iowa Orthop J 2001;21:1–7.
2. Hunter DJ, Felson DT. Osteoarthritis. BMJ 2006 Mar 18;332(7542): 639–42.
3. Cicuttini F, Wluka A, Hankin J, Wang Y. Longitudinal study of the relationship between knee angle and tibiofemoral cartilage volume in subjects with knee osteoarthritis. Rheumatology (Oxford) 2004 Mar;43(3): 321–4.
4. Sharma L. Local mechanical factors in the natural history of knee osteoarthritis. Malalignment and joint laxity. In: Brandt KD, Doherty M, Lohmander LS, Eds. Osteoarthritis. 2nd edn. Oxford: Oxford University Press; 2003:177–83.
5. Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. JAMA 2001 Jul 11;286(2):188–95.
6. Cooke TD, Scudamore A, Greer W. Varus knee osteoarthritis: whence the varus? J Rheumatol 2003 Dec;30(12):2521–3.
7. Zhai G, Ding C, Cicuttini F, Jones G. A longitudinal study of the association between knee alignment and change in cartilage volume and chondral defects in a largely non-osteoarthritic population. J Rheumatol 2007 Jan;34(1):181–6.
8. Swank ML, Lehnert IE. Orthopedic personnel roles in the OR for computer assisted total knee arthroplasty. AORN J 2005 Oct;82(4): 631–43.
9. Iorio R, Healy WL, Richards JA. Comparison of the hospital cost of primary and revision total knee arthroplasty after cost containment. Orthopedics 1999 Feb;22(2):195–9.
10. Moreland JR, Bassett LW, Hanker GJ. Radiographic analysis of the axial alignment of the lower extremity. J.Bone Joint Surg Am 1987 Jun; 69(5):745–9.
11. Kraus VB, Vail TP, Worrell T, McDaniel G. A comparative assessment of alignment angle of the knee by radiographic and physical examination methods. Arthritis Rheum 2005 Jun;52(6):1730–5.
12. Beattie KA, Boulos P, Pui M, O'Neill J, Inglis D, Webber CE, *et al.* Abnormalities identified in the knees of asymptomatic volunteers using peripheral magnetic resonance imaging. Osteoarthritis Cartilage 2005 Mar;13(3):181–6.
13. Boulos P, Beattie KA, Gordon C, Jurriaans E, Kumbhare D, Kaminska E, *et al.* Minimum medial joint space width reproducibility using an automated technique. J Rheumatol 2002;29(7):1566.
14. Peterfy C, Li J, Zaim S, Duryea J, Lynch J, Miaux Y, *et al.* Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test–retest reproducibility. Skeletal Radiol 2003 Mar;32(3):128–32.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8; 1(8476):307–10.
16. Goker B, Block JA. Improved precision in quantifying knee alignment angle. Clin Orthop Relat Res 2007 May;458:145–9.
17. Gluer CC, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. Osteoporos Int 1995;5(4):262–70.
18. Wang Y, Wluka AE, Davis S, Cicuttini FM. Factors affecting tibial plateau expansion in healthy women over 2.5 years: a longitudinal study. Osteoarthritis Cartilage 2006 Dec;14(12):1258–64.
19. Sharma L, Lou C, Cahue S, Dunlop DD. The mechanism of the effect of obesity in knee osteoarthritis: the mediating role of malalignment. Arthritis Rheum 2000 Mar;43(3):568–75.
20. Sailer J, Scharitzer M, Peloschek P, Giurea A, Imhof H, Grampp S. Quantification of axial alignment of the lower extremity on conventional and digital total leg radiographs. Eur Radiol 2005 Jan;15(1): 170–3.
21. Takahashi T, Yamanaka N, Komatsu M, Ogawa Y, Yoshida S, Yamamoto H. A new computer-assisted method for measuring the tibio-femoral angle in patients with osteoarthritis of the knee. Osteoarthritis Cartilage 2004 Mar;12(3):256–9.
22. Sanfridsson J, Ryd L, Eklund K, Kouvaras Y, Jonsson K. Angular configuration of the knee. Comparison of conventional measurements and the QUESTOR precision radiography system. Acta Radiol 1996 Sep;37(5):633–8.
23. Kothari M, Guermazi A, von Ingersleben G, Miaux Y, Sieffert M, Block JE, *et al.* Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis. Eur Radiol 2004 Sep;14(9):1568–73.
24. Botha-Scheepers S, Kloppenburg M, Kroon HM, Hellio Le Graverand MP, Breedveld FC, Ravaud P, *et al.* Fixed-flexion knee radiography: the sensitivity to detect knee joint space narrowing in osteoarthritis. Osteoarthritis Cartilage 2007 Mar;15(3):350–3.
25. Nevitt MC, Peterfy C, Guermazi A, Felson DT, Duryea J, Woodworth T, *et al.* Longitudinal performance evaluation and validation of fixed-flexion radiography of the knee for detection of joint space loss. Arthritis Rheum 2007 May;56(5):1512–20.
26. Ward RJ, Buckland-Wright JC, Wolfe F. Relationships between tibial rim alignment and joint space width measurement reproducibility in non-fluoroscopic radiographs of osteoarthritic knees. Osteoarthritis Cartilage 2005 Nov;13(11):945–52.
27. Cooke TD. Static knee alignment and its association with radiographic knee osteoarthritis. Osteoarthritis Cartilage 2007 Jul;15(7):844–5.
28. Kraus VB. Response to letter: static knee alignment and its association with radiographic knee osteoarthritis. Osteoarthritis Cartilage 2007 Jul;15(7):846–7.